

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2002, 13:02:34 : Search time 1593 Seconds  
(without alignments)  
3304.339 Million cell updates/sec

Title: US-09-818-954A-2

Perfect score: 390

Sequence: 1 atgaagctgcgtctctctc.....ccacgagctgtgagaccatc 390

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues 27472414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estlin:\*

4: em\_estlmu:\*

5: em\_estlov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hlc:\*

9: gb\_estl:\*

10: gb\_estl2:\*

11: gb\_hlc:\*

12: gb\_gss:\*

13: em\_gss\_hum:\*

14: em\_gss\_inv:\*

15: em\_gss\_pln:\*

16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	286.2	73.4	680	9	BB624070 BB624070
2	188	48.2	405	12	AQ495547 HS-5211.B
3	176.6	45.3	560	10	BM262079 d4g40d11.B
4	171.2	43.9	617	9	BE131943 db40d09.Y
5	168.2	43.1	635	10	BM262389 d4g40d11.B
6	72.6	18.6	928	12	CNS0201E
7	57.8	14.8	285	12	BH327754 CH230-117
8	53	13.6	1395	11	AK017593 Mus muscu
9	50.8	13.0	592	12	BH279786 CH230-108
10	49.8	12.8	423	10	BG379148 UI-R-BT1-
11	49.8	12.8	497	9	BE102709 UI-R-BT1-
12	49.4	12.7	434	10	BI033915 OY2-NN200
13	48.8	12.5	606	9	AV754501 AV754501
14	48.2	12.4	372	10	BF398865 UI-R-CAL-
15	48	12.3	452	9	AV746068 AV746068
16	47.8	12.3	268	12	AZ871616 2M0184002
17	47.8	12.3	522	9	AV173269 AV173269

18	47.8	12.3	523	9	BB618381
19	47.8	12.3	546	10	BF074571
20	47.8	12.3	584	10	BM224915
21	47.4	12.2	584	10	BM224971
22	47.2	12.1	453	9	A1051683
23	47.2	12.1	560	9	AV752236
24	47.2	12.1	678	9	AV751658
25	46.4	11.9	328	10	BF079065
26	46.4	11.9	379	10	BF442180
27	46.4	11.9	409	10	BE749737
28	46.4	11.9	414	10	BF075383
29	46.4	11.9	476	10	BF442020
30	46.4	11.9	482	10	BF080835
31	46.4	11.9	515	10	BI343866
32	46.4	11.9	516	10	BI345127
33	46.4	11.9	516	10	BF702350
34	46.4	11.9	527	10	BI339266
35	46.4	11.9	625	9	AV753384
36	45.8	11.7	634	9	AV745740
37	45.4	11.6	481	9	A1189453
38	45.4	11.6	525	9	A1148695
39	45	11.5	544	9	A1148647
40	44.8	11.5	519	9	A1149163
41	44.6	11.4	447	9	A1291937
42	44	11.3	433	9	AV748886
43	44	11.3	487	10	BF190158
44	44	11.3	539	9	A1188964
45	43.8	11.2	446	9	A1276216

## ALIGNMENTS

RESULT 1

BB624070

LOCUS

DEFINITION

musculus CDNA clone 7530426119 5', mRNA sequence.

ACCSSION

BB624070

VERSION

BB624070.1

KEYWORDS

SOURCE

ORGANISM

house mouse.

musculus

REFERENCE

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Iehli,Y., Ito,M., Kawai,U., Konno,H., Kouda,M., Koye,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wagui,K., Fujiwaka,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.



QY 323 CCTATCCCGGCGCATCCGCTGTGACTGCGAGCGCTCTCCATCCAGCCAGGAGTGTG 382  
 Db 142 CCGATCCCGGCGCATCCGCTGTGACTGCGAGCGCTCTCCATCCAGCCAGGAGTGTG 201  
 QY 383 AGACATC 390  
 Db 202 AGACATC 209

RESULT 3  
 BM262079/c 560 bp mRNA linear EST 18-DEC-2001  
 LOCUS dag40d11.x3 Blackshear/Soares normalized Xenopus egg library  
 DEFINITION Xenopus laevis cDNA clone IMAGE:4783917 3' similar to SW:GTHB\_CYPCA  
 P01235 GONADOTROPIN BETA CHAIN PRECURSOR. [1] ; mRNA sequence.  
 ACCESSION BM262079  
 VERSION BM262079.1 GI:17925119  
 KEYWORDS EST.  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 560)  
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Page,D.,  
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person  
 ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,  
 Waterston,R. and Wilson,R.  
 Washu Xenopus EST project, 1999  
 Unpublished (1999)  
 Contact: Sandy Clifton, Ph.D.  
 Washu Xenopus EST project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Library constructed by Bento Soares and M. Fatima Bonaldo  
 (University of Iowa). DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone distribution: Xenopus clones from this library are available  
 through the I.M.A.G.E. Consortium/LNL at: info@image.lnl.gov  
 Seq primer: -40UP from G1bco  
 High quality sequence stop: 473.  
 Location/Qualifiers  
 1..560  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="IMAGE:4783917"  
 /clone\_lib="Blackshear/Soares normalized Xenopus egg  
 library"  
 /sex="female"  
 /tissue\_type="unfertilized egg"  
 /cell\_type="unfertilized egg"  
 /dev\_stage="unfertilized egg"  
 /lab\_host="DH10B"  
 /note="Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI;  
 polyA-selected mRNA was prepared from unfertilized Xenopus  
 laevis eggs. The library was constructed from unfertilized Xenopus  
 pT73-Pac as described in Bonaldo, M.F., Lennon, G. and  
 Soares, M.B. 'Normalization and subtraction: two  
 approaches to facilitate gene discovery', Genome Research  
 6:791-806, 1996. The first strand synthesis used a  
 NotI-drf18 primer; double stranded cDNAs were ligated to  
 EcoRI adaptors, digested with NotI, and directionally  
 cloned into the NotI and EcoRI-digested pT73-Pac vector.  
 The library contained approximately 7.2 x 10<sup>5</sup>  
 recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 164 a 112 c 120 g 164 t  
 ORIGIN  
 Query Match 45.3%; Score 176.6; DB 10; Length 560;  
 Best Local Similarity 73.0%; Pred. No. 1.e-34;

Matches 227; Conservative 0; Mismatches 84; Indels 0; Gaps 0;  
 QY 76 TCCAGTGGAGACCTGGCGACCTTTGTGGGCTGTGCGCGAGGAGTTACTTCTGGCC 135  
 Db 543 TCCATATATAGCTGAGAGAACGTTGATGGATGCTGTGAGGAGATTCATCATTTTACA 484  
 QY 136 AAGAAGCCAGGCTGAGGCGCTTGCAGTACACAGAGATGCTGGGCTGTGAG 195  
 Db 483 AAGAAGCTGGCTGACAGAGTCTGCTGTGACTACTGATGCTGCTGGGCGCTGTAG 424  
 QY 196 ACCTGGAGAAACCATTTCTGGAACCCCTATATTGAAGCCATCATGAGTGTACC 255  
 Db 423 ACCTGGAGAGAACCATCTCCATGATTCCTGATACATAGAGCCACACAGAGTGTGACT 364  
 QY 256 TACAGCAGACCAACAGAGTACTGTCAAGTGCACCACTGTGCCCGGAGTGCACCC 315  
 Db 363 TACATATGAATTAACGTTACTGTAAACCTGCCAAACTCAGCCAGACATTTGACCA 304  
 QY 316 TTCTACACCTATCCCGTGCATCCGCTGTGACGTGCGGAGGCTGCTGACATGCCACAG 375  
 Db 303 TTCTTTACCTACCCACAGTTGCTCATTAAGATGTGACTGTGACATTTGTTCCACTTTACTACA 244  
 QY 376 GAGTGTGAGAC 386  
 Db 243 GAATGTGAGAC 233

RESULT 4  
 BE131943  
 LOCUS db40d09.y1 Blackshear/Soares normalized Xenopus egg library Xenopus  
 DEFINITION laevis cDNA clone IMAGE:3300977 5' similar to SW:GTHB\_CYPCA P01235  
 GONADOTROPIN BETA CHAIN PRECURSOR. [2] SW:GTHB\_CYPCA ; mRNA  
 sequence.  
 ACCESSION BE131943  
 VERSION BE131943.1 GI:8579306  
 KEYWORDS EST.  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 617)  
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Page,D.,  
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person  
 ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,  
 Waterston,R. and Wilson,R.  
 Washu Xenopus EST project, 1999  
 Unpublished (1999)  
 Contact: Sandy Clifton, Ph.D.  
 Washu Xenopus EST project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Library constructed by Bento Soares and M. Fatima Bonaldo  
 (University of Iowa). DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone distribution: Xenopus clones from this library are available  
 through the I.M.A.G.E. Consortium/LNL at: info@image.lnl.gov  
 Seq primer: -40RP from G1bco  
 High quality sequence stop: 467.  
 Location/Qualifiers  
 1..617  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="IMAGE:3300977"  
 /clone\_lib="Blackshear/Soares normalized Xenopus egg  
 library"  
 /sex="female"  
 /tissue\_type="unfertilized egg"  
 /cell\_type="unfertilized egg"

```

/dev_stage="unfertilized egg"
/lab_host="DH10B"
/note="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subcloning: Two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-drl8 primer: double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT7T3-Pac vector.
The library contained approximately 7.2 x 106
recombinants, with average insert sizes of 1-1.5 kb."

```

Query Match	43.98%	Score 171.2	DB 9	Length 617
Best Local Similarity	67.98%	Pred. No. 2.6e-33		
Matches 239; Conservative	0;	Mismatches 113;	Indels 0;	Gaps 0

QY	30	CATGGCCCTCCCTCCCTTCGTGGCTGGCTATGGCTGTGTCCTGGAGTCCCTCAGTGGGAACCT	89
Db	266	CCTCAGAGTTCTGGTTTTAGCCCTGTGTCACAGACCCTGCGAGGATCCAAATTAAGTCT	325
QY	90	GGCAGACCTTTGTGGGCTGTGCGCGTAGAGGAGTTTACTTCTCTGGCCAAAGCCAGGCTG	149
Db	326	GAGACGCTCATTTGGATGATGTGTGAGAGGAAATTCACATTTCTTAGCAAGAAACCTTGCTG	385
QY	150	CAGGGGCGCTTCGGATCACCACAGATGCGCTCTGGGTGCGCTGTAGACACTGGGAGAAACC	209
Db	386	CAGAGGTCTGTGGTGTGACTACTGATGCTCTCTGGGGCGCGCTGTAGACCTGTGAGAAACC	445
QY	210	CATTCTGGAACCCCCCTATATTTGAAGGCCATCATCAGTGTGTAACTTACACAGACAGCAA	269
Db	446	ATCCCTAGATCCTCCGTAACATAGAAAGCCACACAGAGTGTGCACTTAATATAAATCAA	505
QY	270	ACAGCTACTGTCGAAGTGGCCCACTGGCCGGGAGTGAACCCCTTTCATACACTATCC	329
Db	506	ACTGTTACTGTATTACTGGCCAAACTGACGCCAGACATTGACCCANTCTTTAACTAACCC	565
QY	330	CGTGGCCATCCGCTGTGACTGTGGAGGCTGTGCATCTGACCAACGAGAGTGT	381
Db	566	AGTTGACCTTAAGATGTGACTGTGACATGTGGTCCACTTCCTTACACAAAGTGT	617

RESULT 5  
BM262389

LOCUS  
DEFINITION

BM262389 635 bp mRNA linear EST 18-DEC-2001  
day40d11.y3 Blackshear/Soares normalized Xenopus egg library  
Xenopus laevis cDNA clone IMAGE:4783917 5' similar to SW:GTFB\_CYPCA  
P01235 GNADOTROPIN BETA CHAIN PRECURSOR. [1] ; mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS

BM262389  
BM262389.1 GI:17925429  
EST.

SOURCE  
ORGANISM

African clawed frog.  
Xenopus laevis

REFERENCE

Eukaryota: Metazoa; Chordata; Craniata; Vertebrata: Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
Xenopodinae; Xenopus.

AUTHORS

1 (bases 1 to 635)  
Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Miller,L., Pape,D.,  
Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person,  
B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,  
Waterston,R. and Wilson,R.

TITLE

Washu Xenopus EST project, 1999

JOURNAL

Unpublished (1999)

COMMENT

Other\_ESTs: day40d11.x3  
Contact: Sandy Clifton, Ph.D.  
Washu Xenopus EST project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800

Fax: 314 286 1810  
 Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
 Library constructed by Bento Soares and M. Fatima Bonaldo  
 (University of Iowa). DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone distribution: Xenopus clones from this library are available  
 through the I.M.A.G.E. Consortium/LLNL at: [info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Seed primer: -40RP from Gibco  
 High quality sequence stop: 491.  
 Location/Qualifiers  
 1..635

```

/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:4783917"
/clone_lib="Blackshear/Soares normalized Xenopus egg
library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/notes="Vector: pT73-Pac; Site_1: EcoRI; Site_2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT73-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dn18 primer; double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT73-Pac vector.
The library contained approximately 7.2 x 105
recombinants, with average insert sizes of 1-1.5 kb."

```

Query Match	43.1%	Score 168.2	D8 10	Length 635
Best Local Similarity	68.3%	Pred. No. 1.5e-32		
Matches 233	Conservative 0	Mismatches 108	Indels 0	Gaps 0

Accession	Definition	Sequence
QY 30	CATGAGCCCTTCCTCCCTTTGCGTGGCGATAGGGGTGTCTCCGTGGTCCAGTGGGAGACT	89
Db 295	CCTGACAGCTTCTGGTTTAGCCCTGTGTCTCAGACACCCCTGCAGGATTCAAATATAGTCT	354
QY 90	GCGCACCTTTGTGGGCGTGTGCCGTGAGGAGTTTACTTTCTGGCCAAAGAACCCAGGCTG	145
Db 355	GAGAACGTTTACTTATGGATAGTGTCTGTGAGGGAATTCACATTTCTTAGCAAAAGAACCTGGCTG	414
QY 150	CAGAGGCGCTTGGGATCACACGAGATGCCCTGGTGGGCGTGGTGTGAGACTGGGAGATAACCC	209
Db 415	CAGAGGCTTGGCGTGTGACTACTGATGATCCCTGCTGGGAGCGCTGTGTGAGACCTGGGAGAACCC	474
QY 210	CATTCTGGAAACCCCTTATTTGAAGCCCATATCGATGTGTACTTACCTTCAACGAGACCAA	269
Db 475	ATCCCTAGATCCTCCGCTACATAGAAAGCCACACAGACTGTGCATCTTCAATGAAATCAA	534
QY 270	ACAGGTGACTTCAAGCTGCCCAACTGTGCCCGGAGATGCACCCCTCTTACACTATCC	329
Db 535	ACTGTATTACTGTAAAACTGCACAAACTGCAGCCCAACATTAATCCATCTTCTTACTACCC	594
QY 330	CGTGGCCATCCGCTGTGACTCGGAGAGCCTGTCTCCACTGCCA	370
Db 595	AGTTGCCATTGATGTGACTGTGACATTTGTTCACACTCTA	635

RESULT 6  
CNS0201E/c  
LOCUS  
DEFINITION  
CNS0201E 928 bp DNA linear GSS 14-MAY-2000  
Tetradon nigroviridis genome survey sequence T7 end of clone  
152J13 of library G from Tetradon nigroviridis, genomic survey  
sequence.  
AL206123  
AL206123.1 GI:7864942

**KEYWORDS** GSS; genome survey sequence.  
**SOURCE** Tetraodon nigroviridis.  
**ORGANISM** Tetraodon nigroviridis.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.  
**REFERENCE** 1 (bases 1 to 928)  
 Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Queller, F., Saurin, W., Bernot, A. and Weissenbach, J.  
**TITLE** Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 928)  
 Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brothier, P., Queller, F., Saurin, W. and Weissenbach, J.  
**AUTHORS** Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence  
**TITLE** Unpublished  
**JOURNAL** 3 (bases 1 to 928)  
**REFERENCE** Genoscope.  
**AUTHORS** Direct Submission  
**TITLE** Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases  
**JOURNAL** This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.  
**COMMENT** Location/Qualifiers  
 source  
 1..928  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone\_lib="G"  
 /note="Genoscope sequence ID : COAG152CE07JPL-end : T7"  
**BASE COUNT** 266 a 214 c 238 g 196 t 14 others  
**ORIGIN**

Query Match 18.6%; Score 72.6; DB 12; Length 928;  
 Best Local Similarity 77.0%; Pred. No. 3.3e-08;  
 Matches 87; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

**Qy** 98 TTGGGCTGTGGCGTACAGGAGTTACTCTCCGCCCAAGACCGACGCGAGGCC 157  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
**Db** 251 TTCCGGGCTGGCGCGCGGCGGATTCCTCTGTCGCCACAGCCGCGAGGGTGC 192  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
**Qy** 158 TTCCGATCACCACGATGCTGCGGCTGCTGAGACCTGGAGAAACC 210  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
**Db** 191 TGCCGCTCACCACGAGAGCGCTGTGGGGGCTGCTGCCACACCTGGAGGTAACC 139

**RESULT 7**  
**LOCUS** BH327754 285 bp DNA linear GSS 03-DEC-2001  
**DEFINITION** CH230-117A17.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
**ACCESSION** BH327754  
**VERSION** BH327754  
**KEYWORDS** GSS.  
**SOURCE** Norway rat.  
**ORGANISM** Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
**REFERENCE** 1 (bases 1 to 285)  
**AUTHORS** Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn, A., Gebregorgis, F., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C.M.  
**TITLE** Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
**JOURNAL** Unpublished (1999)  
**COMMENT** Other GSSs: CH230-117A17.TV  
 Contact: Shaying Zhao

Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: [szhac@tigr.org](mailto:szhac@tigr.org)  
 Clones are derived from the rat BAC library CHORI-230  
 (<http://www.chori.org/bacpac/rat30.htm>). For BAC library availability, please contact Pieter de Jong ([pdejong@mail.cho.org](mailto:pdejong@mail.cho.org)).  
 Clones may be purchased from BACPAC Resources  
 ([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html)  
 Plate: 117 row: A column: 17  
 Seq primer: 77  
 Class: BAC ends.  
**FEATURES** Location/Qualifiers  
 source  
 1..285  
 /organism="Rattus norvegicus"  
 /strain="BN/SSNhsd/MCw"  
 /db\_xref="taxon:10116"  
 /clone="CH230-117A17"  
 /clone\_lib="CHORI-230 Segment 1"  
 /sex="Female"  
 /cell\_type="Brain"  
 /note="Vector: pTARBAC2.1; Site.1: EcoRI; Site.2: KcoRI; CHORI-230 Rat (BN/SSNhsd/MCw) BAC library produced by Pieter de Jong"  
**BASE COUNT** 60 a 66 c 83 g 76 t  
**ORIGIN**

Query Match 14.8%; Score 57.8; DB 12; Length 285;  
 Best Local Similarity 96.7%; Pred. No. 0.00013;  
 Matches 59; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

**Qy** 144 AGCTGCAGGGGCTTCGATCACCAGATGCTGCGGGCTGCTGAGACCTGGGA 203  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
**Db** 1 AGCTGCAGGGGATTCGATCACCAGATGCTGCGGGCTGCTGAGACCTGGGA 60

**Qy** 204 G 204  
**Db** 61 G 61

**RESULT 8**  
**LOCUS** AK017593 1595 bp mRNA linear HTC 19-JAN-2002  
**DEFINITION** Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730420N04; follicle stimulating hormone beta, full insert sequence.  
**ACCESSION** AK017593  
**VERSION** AK017593.1 GI:12856911  
**KEYWORDS** HTC; CAP trapper.  
**SOURCE** Mus musculus (strain:C57BL/6J) 8 days embryo cDNA to mRNA, clone\_lib:RIKEN full-length enriched mouse cDNA library  
 clone:5730420N04.  
**ORGANISM** Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE** 1 (sites)  
**AUTHORS** Carninci, P. and Hayashizaki, Y.  
**TITLE** High-efficiency full-length cDNA cloning  
**JOURNAL** Meth. Enzymol. 303, 19-44 (1999)  
**MEDLINE** 99279253  
**PUBMED** 10349636  
**REFERENCE** 2 (sites)  
**AUTHORS** Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
**TITLE** Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
**JOURNAL** Genome Res. 10 (10), 1617-1630 (2000)  
**MEDLINE** 20499374  
**PUBMED** 11042159  
**REFERENCE** 3 (sites)

AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Washiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	Riken integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1737-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4 (sites) The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
AUTHORS	5 (bases 1 to 1595) Aachih,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Butt,C., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanaoka,T., Hara,A., Hayashi,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koye,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schirml,L., Shibata,K., Shidate,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Tagami,M., Yamamura,T., Yamataka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGACGAGCAGATCCAGAGCTCTTTTCTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rct = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGACGAGATTCGACGTATTAAATTAATCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmidscript KS(+) after bulk excision from lambda phage I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B. Location/Qualifiers 1..1595 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGD:MGI:1902932" /db_xref="taxon:10090" /clone_id="5730420N04" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="8 days embryo" 64..456 /gene="Fshb" /name="Fshb" 64..456 /note="data source:MGD, source key:MGI:95582, evidence:ISS follicle stimulating hormone beta putative"

/lab/host/BDH10B (Life Technologies)" [/note=Vector: pU73D-Pac \(Pharmacia\)](#) with a modified polylinker. Site.1: Not I. Site.2: Eco RI. The library U1-R101 is a subtracted library derived from a mixture of the following tissues: hippocampus, thalamus, mid-brain, medulla, corpus striatum, cerebral cortex and testis. For a detailed description of the library from which this

```

/cclone="UI-R-BT1-agg-b1-0-UI"
/cclone.lib="UI-R-BT1"
/dec_stage="adult"
/lav_host="DH10B (Life Technologies)"
/note=Vector: pr73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The library
UI-R-BT1 is a subtracted library derived from a mixture of

```

Site 2: SHI. A mini-library was made by cloning products derived from ORSTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions.

BASE COUNT	107 a	102 c	108 g	117 t
ORIGIN				

Query Match	12.7%	Score 49.4	DB 10	Length 434
Best Local Similarity	54.2%	Pred. No. 0.02		
Matches 123	Conservative 0	Mismatches 101	Indels 3	Gaps 1

OY 155 GCCTTGGGATACCAAGGATGCCGTGCGTTGTAGACCTGGGAGAAACCATTTC 214  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 302 GCATTACGATCAACACCACCTTGCTGCTGCTACTGCTACACCAAGGATCTGGTGTATA 243

OY 155 GCCTTCGGATCACCAGGATGCCCTGTGG66TGCCTTGAGACCTGGGAGAAACCATTC 214  
| | | | | | | | | | | | | | |  
Db 302 GCATTACGATCAACCCACTTTGCTGTCTGTGCTACTCTACACCA66GATCTGGTGTATA 243

Oy 335 CCATCCGCTGTGACTCGGAGCCTGTCTCACCTGCACCACGAGTGT 38  
||| ||| ||| ||| | | | | | | |  
Db 125 CCACCCAGTGTCACCTGTGGCACTGTGACACCGACGACGACTGATTGT 79

OY	335	CCATCCGCTGTGAC	TGCGAGAGCC	CTTCACAT	TGCACAC	CAGCAGAG	TCT	381
D6	125	CCACCCAGTGTCACT	GTGTGCAAGTGT	GCACAGCAG	CAGCAGTAT	TGT		79

OY	335	CCATCCGCTGTGAC	TGCAGCGAGCCTTC	CACATGCCACCACGAGAGT	381
D6	125	CCACCCAGTGCACT	GTTGTGCAACGAGCATGATTGT		79

OY	335	CCATCCGCTGTGAC	TGCGAGAGCC	CTTCACAT	TGCACAC	CAGCAGAG	TCT	381
D6	125	CCACCCAGTGTCACT	GTGTGCAAGTGT	GCACAGCAG	CAGCAGTAT	TGT		79

BASE COUNT	107 a	102 c	108 g	117 t
ORIGIN				

1:

1:

1:

1:

1:

OY	335	CCATCCGCTGTGAC	TGCAGCGAGCCTTC	CACATGCCACCACGAGAGT	381
D6	125	CCACCCAGTGCACT	GTTGTGCAACGAGCATGATTGT		79

OY	335	CCATCCGCTGTGAC	TGCAGCGAGCCTTCAC	TGGCACCCAGCAGTCT	381
D6	125	CCACCACAGTGCA	GTTGTGACACGACGACGACTA	TGTT	79

OY	335	CCATCCGCTGTGAC	TGCAGCGAGCCTTC	CACATGCCACCACGAGAGT	381
D6	125	CCACCCAGTGCACT	GTTGTGCAACGAGCATGATTGT		79



us-09-818-954a-2.rst

Matches	80:	Conservative	0:	Mismatches	52:	Indels	0:	Gaps	0:
QY	250	TTGTAACCTCAACAGACCAACCAAGTAGCTGACTGTAACAGTCCGCCCACTGTGCCCCGGAGATC	309						
Db	207	TGTACTTCCTCAAGGAAGTGTATATGAAACAGAGAGAGTCCCGGCTGTCTCCATCATCA	266						
QY	310	GACCCCTCTCAACACTATACCGGTGGCCATCGCGTGTGACTGGGAGCGCTGCCTCCACTGCC	369						
Db	267	GATTCTCTGTATACCTATTACCAAGTGGCCACCAAGTGTACTGTGGCAAGTGTGACAGCAGAC	326						
QY	370	ACACGACGAGTGT	381						
Db	327	AGCACTGATGTGT	338						
RESULT 14									
BF398865/c									
LOCUS									
DEFINITION									
ACCESSION									
VERSION									
KEYWORDS									
ORGANISM									
SOURCE									
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
MEDLINE									
COMMENT									

Query Match	12.4%	Score 48.2	DB 10	Length 372
Best Local Similarity	54.9%	Pred. No. 0.038		
Matches	95	Conservative	0	Mismatches 78; Indels 0; Gaps 0;
OY	212	TTCTGGAAACCCCCCATATTGTGAAGCCCATTCATCGAGTGTGTACTACACAGACCAAC	271	
Db	332	TACTGCGCAGCGTCATATGGCTCCCGCCCTGACGACGAGTGTGACCTACCGTAGCGCGCT	273	
OY	272	AGGTGACGTGAAGCTGGCCCAACTGTGGCCGGGAGTGGACCCCTTACACACTATCCCG	331	
Db	272	TGCGCTGTGTCCGCTCCCTGGCTGCCACCTGGGTGTAGACCCCATATGCTCTTTCCTG	213	
OY	332	TGGCCATCCGCTGTGACTGGCGAGCCCTGCTCCACTGTCACACGAGAGTGTAG	384	
Db	212	TGGCCCTGAGCGCGCTGTGTGGGCCCTCCGCTTCATGAGCTGTGACTGTGGG	160	
RESULT 15				
LOCUS	AV746068	452 bp	mRNA	linear
DEFINITION	AV746068 NPA Homo sapiens cDNA clone NPADE10 5', mRNA sequence.			
ACCESSION	AV746068			
VERSION	AV746068.1	GI:10865515		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	I (bases 1 to 452)			
TITLE	Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q., Han,Z., Chen,Z., Hu,R. and Chen,J.			
JOURNAL	Homo sapiens NPA library cDNA clones			
COMMENT	Unpublished (2000)			
	Contact: Qinghua Zhang			
	Shanghai Institute of Endocrinology, Rui-Jin Hospital			
	197 Rui-Jin II Road, Shanghai 200025, P. R. China			
	Tel: 86-21-64370045(ex.663332)			
	Fax: 86-21-64743206			
	Email: mbsheims.stn.sh.cn			
	This clone is available at Shanghai Hematology Institute in Shanghai.			
FEATURES	Chinese National Human Genome Center at Shanghai			
	351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.			
source	Location/Qualifiers			
	1..452			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="NPADE10"			
	/clone_lib="NPA"			
	/tissue_type="pituitary"			
	/dev_stage="Adult"			
	/lab_host="SOLR"			
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"			
BASE COUNT	58 a 177 c 122 g 91 t 4 others			
ORIGIN				
Query Match	12.3%	Score 48	DB 9	Length 452
Best Local Similarity	57.1%	Pred. No. 0.045		
Matches	84	Conservative	0	Mismatches 63; Indels 0; Gaps 0;
OY	238	CATCATGAGTCTGACTACACAGACAAACAGGAGCTGTGACACTGCCCAACTGT	297	
Db	166	CCTGAGGTGTGTGACCTACCGGTATGTGGCGCTTCGAGTGCATCCGGCTCCCTGGCTGC	245	
OY	238	GCCCGGAGTGCAGCCCTTCTTACACCTATCCCGTGGCCATCCGCTGTGACTGGAGCC	357	
Db	246	CCGCTGTGNGTGAGCCCGTGTCTCTCCCTGTGAGCTTCAGCTGTGNGAGCC	305	

QY 358 TGCTCCACTGCGACGAGGTGTGAG 384  
||| |  
Db 306 TGCCGCGNAGCCTCTGACTGTGGG 332

Search completed: October 12, 2002, 13:29:29  
Job time : 1599 secs